REPORT DOCUMENTATION PAGE

Form Approved OMB NO. 0704-0188

1. REPORT DATE (DD-MM-YYYY)	2. REPORT TYPE	3. DATES COVERED (From - To)
13-01-2012	Final Report	1-Jun-2007 - 31-May-2011

4. TITLE AND SUBTITLE
Systems Biology of Dehalococcoides: Using Network Inference Modeling to Integrate Omics Datasets Under Varied Conditions

5a. CONTRACT NUMBER
W911NF-07-1-0249

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER
611102

6. AUTHORS
Ruth Richardson

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAMES AND ADDRESSES
Cornell University
Office of Sponsored Programs
Cornell University
Ithaca, NY 14853 -2801

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
U.S. Army Research Office
P.O. Box 12211
Research Triangle Park, NC 27709-2211

10. SPONSOR/MONITOR'S ACRONYM(S)
ARO

11. SPONSOR/MONITOR'S REPORT NUMBER(S)
52345-LS.6

12. DISTRIBUTION AVAILIBILITY STATEMENT
Approved for Public Release; Distribution Unlimited

13. SUPPLEMENTARY NOTES
The views, opinions and/or findings contained in this report are those of the author(s) and should not contrued as an official Department of the Army position, policy or decision, unless so designated by other documentation.

14. ABSTRACT
The overall objectives of this project were to 1. obtain a systems-biology level understanding of gene networks in the Dehalococcoides and 2. develop assays for quantitative biomarkers of chloroethene detoxification steps and rates that could be deployed at sites undergoing in situ bioremediation utilizing this important group of microbes. After collecting genome-wide microarray expression data along with metabolite and chloroethene data, under a wide range of conditions (n = 53), we employed Bayesian inference algorithms to reconstruct the gene-gene and

15. SUBJECT TERMS
Dehalococcoides, Systems biology, microarrays, dehalorespiration, transcriptional regulation, reductive dehalogenases, bioremediation, proteomics, multiple reaction monitoring

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	15. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			Ruth Richardson
UU	UU	UU	UU		19b. TELEPHONE NUMBER: 607-255-3233

Standard Form 298 (Rev 8/98)
Prescribed by ANSI Std. Z39.18

Report Title

Systems Biology of Dehalococcoides: Using Network Inference Modeling to Integrate Omics Datasets Under Varied Conditions

ABSTRACT

The overall objectives of this project were to 1. obtain a systems-biology level understanding of gene networks in the Dehalococcoides and 2. develop assays for quantitative biomarkers of chloroethene detoxification steps and rates that could be deployed at sites undergoing in situ bioremediation utilizing this important group of microbes. After collecting genome-wide microarray expression data along with metabolite and chloroethene data, under a wide range of conditions (n = 53), we employed Bayesian inference algorithms to reconstruct the gene-gene and gene-metabolite subnetworks that are most supported by the expression data. These networks then lead to the discovery of robust biomarker candidates of respiration as well as of stress. Quantitative assays for these biomarkers were then developped and applied to examine the power of protein and/or RNA biomarker levels to serve as estimators of bulk culture respiration rates.

Enter List of papers submitted or published that acknowledge ARO support from the start of the project to the date of this printing. List the papers, including journal references, in the following categories:

(a) Papers published in peer-reviewed journals (N/A for none)

Number of Papers published in peer-reviewed journals: 2.00

(b) Papers published in non-peer-reviewed journals (N/A for none)

Received	Paper

TOTAL:

Number of Papers published in non peer-reviewed journals: 0.00

(c) Presentations

Number of Presentations: 0.00

Non Peer-Reviewed Conference Proceeding publications (other than abstracts):

Received	Paper

TOTAL:

Number of Non Peer-Reviewed Conference Proceeding publications (other than abstracts): 0

Peer-Reviewed Conference Proceeding publications (other than abstracts):

Received	Paper

TOTAL:

Number of Peer-Reviewed Conference Proceeding publications (other than abstracts): 0

(d) Manuscripts

Received	Paper
2012/01/13 14	3 Annette Rowe, Brendan Lazar, Rober Morris, Ruth Richardson. Characterization of the Community Structure of a Dechlorinating MixedCulture and Comparisons of Gene Expression in Planktonic andBiofloc-Associated “Dehalococcoides” and Methanospirillum Species, Applied and Environmental Microbiology (02 2008)
2012/01/13 14	2 Jeffrey Werner, Brian Rahm, A. Celeste Ptak, Sheng Zhang, Ruth Richardson. Absolute quantification of Dehalococcoides proteins:enzyme bioindicators of chlorinated ethenedehalorespiration, Environmental Microbiology (06 2008)
2012/01/13 14	1 Annette Rowe, Gretchen Heavner, Cresten Mansfeldt, Jeffrey Werner, Ruth Richardson. Relating Chloroethene Respiration in Dehalococcoides to Protein and mRNABiomarkers: Implications for Protein Production and Enzyme-Specific RateParameters, International Journal of Microbial Ecology Journal (ISME Journal) (01 2012)

TOTAL: **3**

Number of Manuscripts: 0.00

Books

Received	Paper

TOTAL:

Patents Submitted

Patents Awarded

Awards

Graduate Students

NAME	PERCENT_SUPPORTED
Brian Rahm	0.13
Annette Rowe	0.25
FTE Equivalent:	**0.38**
Total Number:	**2**

Names of Post Doctorates

NAME	PERCENT_SUPPORTED
Jeffrey Werner	0.50
Brian Rahm	0.05
FTE Equivalent:	**0.55**
Total Number:	**2**

Names of Faculty Supported

NAME	PERCENT_SUPPORTED	National Academy Member
Ruth Richardson	0.00	No
Steven Zinder	0.00	No
James Gossett	0.00	No
FTE Equivalent:	**0.00**	
Total Number:	**3**	

Names of Under Graduate students supported

NAME	PERCENT_SUPPORTED	
Ryan Walter	0.02	
JuKhuan Pan	0.00	Civil Engineering
Sage Hellerstedt	0.00	Civil Engineering
FTE Equivalent:	**0.02**	
Total Number:	**3**	

Student Metrics

This section only applies to graduating undergraduates supported by this agreement in this reporting period

The number of undergraduates funded by this agreement who graduated during this period:	3.00
The number of undergraduates funded by this agreement who graduated during this period with a degree in science, mathematics, engineering, or technology fields:.....	3.00
The number of undergraduates funded by your agreement who graduated during this period and will continue to pursue a graduate or Ph.D. degree in science, mathematics, engineering, or technology fields:.....	2.00
Number of graduating undergraduates who achieved a 3.5 GPA to 4.0 (4.0 max scale):	2.00
Number of graduating undergraduates funded by a DoD funded Center of Excellence grant for Education, Research and Engineering:	0.00
The number of undergraduates funded by your agreement who graduated during this period and intend to work for the Department of Defense	0.00
The number of undergraduates funded by your agreement who graduated during this period and will receive scholarships or fellowships for further studies in science, mathematics, engineering or technology fields:	1.00

Names of Personnel receiving masters degrees

NAME
Total Number:

Names of personnel receiving PHDs

NAME
Brain Rahm, 2008
Annette Rowe, 2011
Total Number: 2

Names of other research staff

NAME	PERCENT_SUPPORTED
FTE Equivalent:	
Total Number:	

Sub Contractors (DD882)

1 a. Gene Network Sciences

1 b. 58 Chalres Street

Cambridge MA 02141

Sub Contractor Numbers (c):

Patent Clause Number (d-1):

Patent Date (d-2):

Work Description (e): Network inference from microarray datasets under a variety of conditions.

Sub Contract Award Date (f-1):

Sub Contract Est Completion Date(f-2):

Inventions (DD882)

Scientific Progress

See attachment

Technology Transfer

PROJECT TITLE: SYSTEMS BIOLOGY OF DEHALOCOCCOIDES: USING NETWORK INFERENCE MODELING TO INTEGRATE OMICS DATASETS UNDER VARIED CONDITIONS

DoD Award: W911NF-07-1-0249

ABSTRACT

The overall aims of this project were to 1. obtain a systems-biology level understanding of gene networks in the Dehalococcoides and 2. develop assays for quantitative biomarkers of chloroethene detoxification steps and rates that could be deployed at sites undergoing in situ bioremediation utilizing this important group of microbes. After collecting genome-wide microarray expression data along with metabolite and chloroethene data, under a wide range of conditions (n = 53), we employed Bayesian inference algorithms to reconstruct the gene-gene and gene-metabolite subnetworks that are most supported by the expression data. These networks then led to the discovery of robust biomarker candidates of respiration (key dehalogenases and the core hydrogenase) as well as of stress (e.g chaperones, heat shock proteins, superoxide dismutase). Quantitative assays for these biomarkers were then developed and applied to examine the power of protein and/or RNA biomarker levels to serve as estimators of bulk culture respiration rates. Collectively the work is the most comprehensive picture of Dehalococcoides' systems biology yet available. On the applied side, RNA biomarkers were responsive to respiration rate down to 0.01x the populations' maximum – thereby demonstrating a wide range over which they are informative. These RNA-respiration trends were also tested in an additional strain of Dehalococcoides (KB1 – a commercially sold bioaugmentation culture that has been deployed at dozens of sites across the country) and results for RNA from the hydrogenase enzyme "Hup" were remarkable consistent across the organisms in trend and absolute abundance (in transcripts per cell). Application of the mass spectrometry-based protein biomarker quantification method we developed enabled calculation of enzyme kinetic rate parameters *in vivo* for two key reductive dehalogenase enzymes: TCE-Reductive dehalogenase and PCE-Reductive dehalogenase.

Statement of the problem studied

Organisms in the group "*Dehalococcoides*" are remarkable for their ability to respire a wide range of chlorinated organic compounds. Application of DHC physiology at field sites has

grown widespread for the treatment of groundwater chlorinated solvent plumes (PCE and TCE) and there is growing interest in extending application to bioremediation of PCB, chlorinated benzenes, and chlorinated phenols. However, *in situ* bioremediation requires effective documentation to government agencies. mRNA/protein biomarkers that are specific to detoxification processes are being explored for exactly this purpose. Additionally, at sites where remediation has stalled or slowed to unacceptable rates, it would be useful to monitor stress response biomarkers to aid site managers in appropriate corrective action.

Another major hurdle toward better understanding of the Dehalococcoides' biology is the fact that no genetic system has yet succeeded – thereby precluding traditional strategies for functional annotation of the organisms' ~1600 genes. Another strategy must therefore be employed to suggest gene function. In this work we took the approach of combining g genome-wide gene expression profiling with Bayesian inference algorithms to generate data-driven models of gene networks and gene-metabolite interactions.

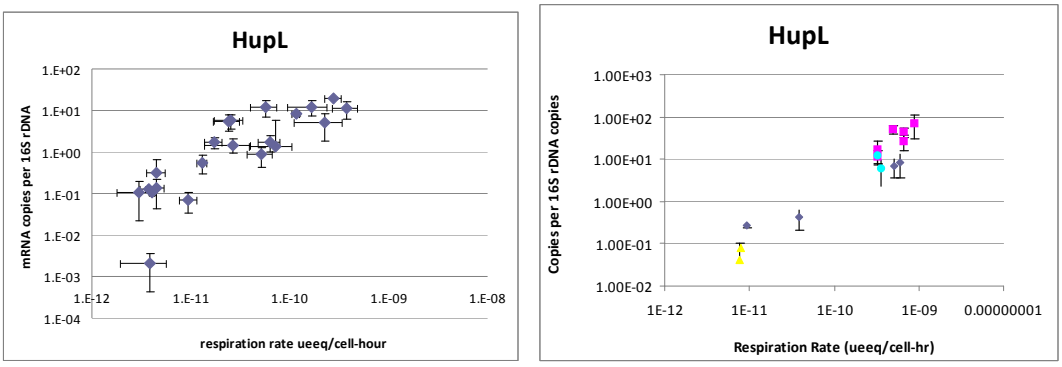
Summary of the most important results

Detailed results for this work are provided in the mansucripts, technical reports, and PhD dissertations uploaded as part of this final report. Here we quickly touch upon some of the key findings vis-à-vis questions posed in the original proposal. These questions, listed below, are addressed in the subsequent paragraphs.

1. Are biomarkers expressed in other cultured/field strains of *Dehalococcoides*? How do candidate biomarker levels correlate with respiration rate over a wider range of growth conditions?
2. What networks of RDases emerge in cultures grown on different substrates? Are there specific transcriptional regulators with expression tied to individual or groups of RDases?
3. Are individual RDases co-regulated with other elements of the proposed electron transport chain (e.g Hup)?
4. Which gene networks correlate with the presence of other community members?
5. Which biomarkers are indicative of DHC stress?

1. Are biomarkers expressed in other cultured/field strains of *Dehalococcoides*? How do candidate biomarker levels correlate with respiration rate over a wider range of growth conditions?

Shotgun proteomics were conducted in an additional three Dehalococcoides-containing cultures from three different labs around the world. In all cases the same key hydrogenase was confirmed at the protein level ("Hup"), while reductive dehalogenases varied from culture to culture. Quantitative reverse transcription PCR (qRTPCR) in two cultures showed remarkable similarity in both the trends between mRNA levels and respiration rates as well as the absolute levels (transcripts per cell). The vast majority of this work is presented in an uploaded manuscript (Rowe et al, title: "Relating Chloroethene Respiration in *Dehalococcoides* to Protein and mRNA Biomarkers: Implications for Protein Production and Enzyme-Specific Rate Parameters") and the comparison between the two cultures in the most promising respiration biomarker (Hup) is given in the following figures (*Dehalococcoides* strain 195 in the left panel and *Dehalococcoides* KB1 1n the right panel). In both cases, the trends are linear on the log-log plots (power law relationship) over two orders of magnitude of per-cell respiration rates.



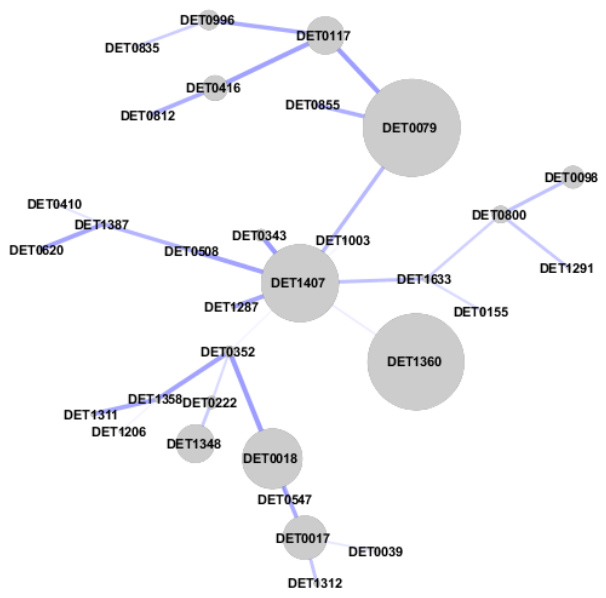
2. What networks of RDases emerge in cultures grown on different substrates? Are there specific transcriptional regulators with expression tied to individual or groups of RDases?

The complete dataset altered both the electron donor and acceptor rate and type fed to the Donna II mixed community culture at a continuous rate as mentioned previously. Within the chlorinated ethene family, the previously monitored reductive dehalogenases of tceA (DET0079), pceA (DET0318), DET1559, and DET1545 were detected –but none were specific to a particular compound, rather they were expressed whenever any chlorinated ethene was ed. Additionally, the reductive dehalogenases DET0180 and DET0174 displayed statistically confident

representation. DET0162, the reductive dehalogenase with the known and reported point mutation, also displayed a significant intensity. The remaining RDases did not distinguish themselves from the background intensity of the microarrays preformed when average intensities across all 53 experiments were averaged. However, in individual experiments some of the "minor" RDases show strong up regulation well above background intensity (e.g. DET1171's specific response to addition of chlorphenols as presented in the manuscript by Mansfeldt et al.2012; NOTE: as this manscript has not yet been submitted is was listed as a Technical Report). DET1171's very specific expression when grown on chlorophenols suggest it has a unique function against this compound class.

In the Gene network analysis, several distinct networks emerge with the RDases mentioned above. The transcript for tceA (DET0079) resides in a gene network (Figure #) with some other expressed transcripts such as the putative S-layer Protein DET1407 and a putative transcriptional regulator (DET0117).

The DET1559 reductive dehalogenase transcript, its associated anchoring protein (DET1558), and its histidine-kinase response regulator (HK/RR) pair (DET1560,1561) all are linked in a



DET0017 DET0017 hypothetical protein (NCBI)
DET0018 DET0018 hypothetical protein (NCBI)
DET0039 DET0039 hypothetical protein (NCBI)
DET0079 DET0079 trichloroethene reductive dehalogenase (NCBI)
DET0098 DET0098 hypothetical protein (NCBI)
DET0117 DET0117 transcriptional regulator, putative (NCBI)
DET0155 DET0155 resolvase domain protein (NCBI)
DET0222 DET0222 serine protease inhibitor family protein (NCBI)
DET0343 DET0343 cell division protein FtsZ (NCBI)
DET0352 DET0352 hypothetical protein (NCBI)
DET0410 DET0410 nodulation ATP-binding protein I (NCBI)
DET0416 DET0416 hypothetical protein (NCBI)
DET0508 DET0508 hypothetical protein (NCBI)
DET0547 DET0547 HDIG domain protein (NCBI)
DET0620 DET0620 hypothetical protein (NCBI)
DET0800 DET0800 hypothetical protein (NCBI)
DET0812 DET0812 PAP2 family protein (NCBI)
DET0855 DET0855 hypothetical protein (NCBI)
DET0996 DET0996 ribosomal protein L33 (NCBI)
DET1003 DET1003 hypothetical protein (NCBI)
DET1206 DET1206 homoserine dehydrogenase (NCBI)
DET1287 DET1287 transcriptional regulator, MerR family (NCBI)
DET1291 DET1291 competence/damage-inducible protein CinA C-terminal domain protein (NCBI)
DET1311 DET1311 hypothetical protein (NCBI)
DET1312 DET1312 hypothetical protein (NCBI)
DET1348 DET1348 RNA polymerase sigma-70 factor, ECF subfamily (NCBI)
DET1358 DET1358 type IV pilus assembly family protein (NCBI)
DET1360 DET1360 general secretion family protein (NCBI)
DET1387 DET1387 hypothetical protein (NCBI)
DET1407 DET1407 BNR/Asp-box repeat domain protein (NCBI)
DET1633 DET1633 aspartate kinase, monofunctional class (NCBI)

single network. This supports both the predicted operon structure of this reductive dehalogenase as well as the relationship the RDase has with the histidine-kinase receptor. However, histidine-kinase receptors regulate the expression of the gene transcript by their phosphoralization state and therefore the linking of the expression values must be analyzed with caution.

3. Are individual RDases co-regulated with other elements of the proposed electron transport chain (e.g Hup)?

Surprisingly, the only RDases that clusters with Hup is one of the "minor" RDases – DET1545. Interestingly, though, of all the genes in the genome, Hup's strongest association was with a putative NADH/ubiquinone oxidoreductase ("Nuo") (edge score between Hup & Nuo = 0.849). This has led to the hypothesis that electrons stripped from H2 are first passed to Nuo as the first step in the electron transport chain in these organisms. Biochemical assays with purified enzymes would be needed to confirm this proposed interaction.

4. Which gene networks correlate with the presence of other community members?

In addition to the DET genes, the microarray also had probes for other community members' 16S rRNA genes. Although the vast majority of edges in the REFS Gene Network model, there were several cases of DET genes' interacting strongly with other community members 16S rRNA intensity. These include several Gram positive organisms. The Gram positives are now thought to be the major butyrate fermenters in the culture and *Dehalococcoides* is dependent upon them for the H2 gas utilized as the sole electron donor.

5. Which biomarkers are indicative of DHC stress?

Cultures were exposed to the following stresses in these studies: PCE starvation, PCE solvent effect, and O2 stress. A list of the Clusters of Orthologous Genes (COGs) categories in *D. ethenogenes was obtained using the* Joint Genome Institute (http://img.jgi.doe.gov). The stress response genes were predicted to be annotated with the COGs of "posttranslational modification, protein turnover, chaperones" and "replication, recombination and repair". A total of 141 potential stress response genes were identified in the genome, with 57 from the "posttranslational modification, protein turnover, chaperones" COG and 84 from the "replication, recombination and repair" COG. This number is corrected for the gene products which have been annotated twice with different locus tags. Of these many were expressed highly at the RNA level. However, most were highly expressed in control cultures that were actively and happily respiring. Using a combination of the microarray and shotgun proteomics data, a surpising gene was identified as specifically responding to Oxygen stress: DET1132, which codes for a protein involved in glutamine biosynthesis. Why this core metabolic protein would be beneficial to handling the oxidative stress is as yet unclear.